Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Zero-Touch Service Management: AI-Native BSS/OSS Evolution In Edge-Cloud Environments

Balakumar Ravindranath Kunthu

kunthu_balakumar@yahoo.com
orcid=0009-0002-9220-5750

The convergence of artificial intelligence, edge computing, and cloud-native architectures fundamentally transforms telecommunications Business Support Systems and Operations Support Systems, enabling autonomous zero-touch service management paradigms. This comprehensive analysis examines architectural frameworks, performance metrics, and economic impacts of AI-native BSS/OSS platforms designed for distributed edge-cloud environments requiring sub-millisecond responsiveness and autonomous operation. The zero-touch provisioning market demonstrates robust growth from \$3.76 billion in 2025 toward projected \$10.02 billion by 2035 at 10.3% compound annual growth rate, while cloud OSS/BSS segments expand from \$44.21 billion to \$56.85 billion by 2030. Quantitative findings reveal autonomous network implementations deliver average annual benefits of \$800 million per communications service provider, comprising \$300 million CapEx savings, \$350 million OpEx reductions, and \$144 million revenue uplift. Zero-touch operations achieve 96% service provisioning acceleration (24 hours to 0.5 hours), 93% incident resolution improvement (180 to 12 minutes), and 65% operational cost reduction compared to manual processes. Edge computing integration enables 60-80% latency reduction critical for ultra-reliable low-latency communications, while AI-driven fault detection accuracy reaches 95% versus 65% for manual approaches. Comprehensive cost-benefit analysis indicates 213% return on investment over three-year implementation cycles, with network availability improvements from 99.0% to 99.9% and energy efficiency gains of 42% through autonomous optimization.

Keywords
- Zero-Touch Automation
- AI-Native BSS/OSS
- Edge Computing
- Autonomous Networks
- Self-Healing Systems
- Cloud-Native Architecture
- Service Orchestration
- Predictive Maintenance
- Intent-Based Operations
- Telecommunications

1. Introduction

Many factors are contributing to the exponentially increasing complexity of telecommunications networks, such as the implementation of 5G, network function virtualization, software-defined networking, the proliferation of the Internet of Things, and the distribution of edge computing. Traditional Operations Support Systems and Business Support Systems are structured to cater to centralized control planes and human-mediated workflows, and therefore, they cannot effectively deal with distributed, dynamically reconfigurable infrastructures spanning cloud data centers, regional edge facilities, and local access points. Zero-touch service management, which entails the use of AI-driven closed-loop systems that not only configure and optimize their operations but also permit or disallow human intervention without the need for human supervision, is indicative of the paradigm shift towards autonomous operations.

Quantitative performance metrics serve to corroborate business cases, with examples such as autonomous network implementations leading to an average annual benefit of $800 million per communications service provider, which is composed of a $300 million reduction in CapEx through optimized infrastructure planning, a $350 million reduction in OpEx through operational automation, and a $144 million increase in revenue from new service models enabled by network programmability.

2. Architectural Foundations and Evolution

2.1 Legacy Limitations and Transformation Imperatives

The evolution of the telecom industry gave rise to traditional BSS/OSS platforms as networks stabilized in terms of topologies, traffic patterns became more predictable, and the service catalogs were limited to voice, messaging, and basic data connectivity. Identifiable functions were carried out by monopolistic applications such as inventory tracking, equipment alarm detection and repair monitoring, and order management processing. These functions were performed through highly manual workflows which were extended system boundaries. The said architectures were based on the premises of centralized network management, batch processing sufficiency, and human decision-making for non-routine situations. Newly developed telecommunications environments are in direct conflict with these environment assumptions. Networks virtualization leads to the instantiation of software-based network elements in a distributed cloud infrastructure thus creating temporary topologies that are beyond human operators' tracking capabilities. This is network function virtualization. Software-defined networking is a tool for on-demand network behavior reprogrammability which entails the separation of control planes from data plane operations. The edge computing, which is the transfer of processing to thousands of micro-data centers, the management becomes more complex alongside the need for sub-second response latencies which cannot be mediated by humans (Bello et al., 2024).

2.2 Cloud-Native Architecture Principles

Cloud-native BSS/OSS implementations use asynchronous messaging and RESTful APIs to decompose monolithic applications into containerized microservices. The change in architecture achieves the scale of the components independently according to the workload, continuous integration and deployment facilitate by which the feature velocity is increased and the technology stack evolution is facilitated by the way of the platform being replaced by a wholesale is not necessary. Container orchestration platforms like Kubernetes offer uniform "deployment and lifecycle" capabilities across the different cloud infrastructures (public hyperscale platforms, private telecommunications operator clouds, and edge computing facilities) because distributed microservices can bring the information in real-time from the network events, customer interactions, and business process triggers. Stream processing engines are always examining event flows, recognizing patterns, finding anomalies, and enabling automated workflows without the necessity of batch processing delays typical for legacy systems. A real-time processing platform is imperative for zero-touch automation which requires decision latencies of the order of a millisecond and the immediate execution of the action to maintain operational consistency.

2.3 AI-Native Integration and Autonomous Capabilities

Autonomous Capabilities. Machine learning is a part of the AI-based system's design, where artificial intelligence is considered one of the basic capabilities instead of an additional feature. Data models are tuned for ML workflows by the use of feature stores that unify reusable transformations, model registries, and inference. Event schemas provide detailed contextual information for model training and inference, while API designs facilitate interaction with AI services that perform tasks such as anomaly detection/alert maintenance, intelligent orchestration (using natural language processing), and so on. TM Forum's Autonomous Network framework defines different maturity levels that enable the autonomous capabilities to improve. Level 0 shows the degree of manual work for all tasks. The assisted intelligence of Level 1 supports human operators by providing recommendations. At Level 2 some workflows are only partially automated, but still, human supervision is not required. At Level 3, systems operate beyond direct control in making decisions within certain limits, thus they have conditional autonomy. High autonomy is reflected at Level 4, where systems are capable of functioning alone in a majority of scenarios. Level 5 is full autonomy where networks, being self-regulating, need no human intervention apart from policy making and strategic oversight (Gharbaoui et al., 2022).

3. Zero-Touch Automation Technologies

3.1 Intelligent Orchestration and Intent-Based Operations

Intent-based networking scenarios make it possible to define high-level business goals like getting the best network slice for communication between autonomous vehicles with a latency of 1ms and a reliability of 99.999%. Hence, autonomous systems convert the intentions into technical realizations, select the best resource utilization, and adjust to the altered situations while still achieving the objectives. To comprehend the dependence of resources and constraints on the intent, natural language processing engines interpret the policy specifications; graph neural networks demonstrate resource dependences and constraints; and reinforcement learning algorithms determine the best implementation strategy by trial and error. Service orchestration engines have access to multi-domain and multi-vendor scenarios such as radio access networks, core networks and edge computing facilities, or cloud data centers. The embedded software frameworks provide system coordination between various services. Template-based workflows help to create the high-level process sequences, while the execution can be changed according to context and resource availability. The orchestration platforms' integration with network controllers is supported by standard APIs like ETSI NFV MANO interfaces, TM Forum Open API's architectures, and ONAP architecture.

3.2 Predictive Analytics and Self-Healing Mechanisms

Healing Mechanisms. By using supervised learning algorithms trained on historical failure data, predictive maintenance systems can forecast components, component failures, equipment degradations and performance anomalies that could occur before service-affecting incidents. Telemetry data can be analyzed using time-series analysis techniques, including Prophet algorithms and Long Short-Term Memory networks, to model temporal patterns and detect impediments that indicate the presence or absence of failures. Optimized preventive maintenance scheduling during planned maintenance windows is made possible by survival analysis models that compute probability distributions for component lifespans, eliminating the need for emergency interventions during outages. Identify faults, identify root causes, carry out remediation actions, and verify system effectiveness through closed-loop automation, which employs self-healing capabilities. Despite the existence of anomaly detection algorithms, network telemetry and customer experience metrics are not always used to identify anomalies. To identify fault sources in intricate distributed systems, root cause analysis engines employ correlation techniques, dependency mapping and causal inference.etna (models). Remediation libraries contain proven resolution procedures, while policy engines encode business rules and safety limitations governing automated actions. Validation mechanisms guarantee restoration through automated testing and continuous monitoring, with human operators taking over when automated interventions fail (Habibi et al., 2023).

3.3 Edge Intelligence and Distributed Decision Making

Position processing capabilities in edge computing architectures are, in most cases, moved closer to data sources and service consumption points. This substantially lowers application latency that requires real-time and also backhaul bandwidth usage. Additionally, regulatory requirements for data sovereignty and privacy concerns are satisfied through the use of AI models deployed at the edge, which can determine local telemetry without sending raw data to centralized cloud platforms. Federated learning methods, in this case, help to improve the global models' accuracy trained on scattered edge datasets by giving access to more local datasets and also by cutting down the data sending volume. The benefit of distributed decision-making frameworks is that the edge nodes can function without any interference during the times when there is no connection to the centralized control planes, thus they can guarantee the serving continuity over network partitions. Also, local policy engines that comprise business rules, service level agreement requirements, and operational constraints facilitate independent decision-making.

4. Market Dynamics and Economic Impact

4.1 Zero-Touch Automation Market Growth

The zero-touch provisioning market demonstrates a significant expansion that strongly reflects the industry's recognition of automation as a strategic imperative. The market valuation stood at $3.76 billion in 2025 with short-term projections pointing to an increase up to $8.1 billion by 2032 and long-term forecasts reaching $10.02 billion by 2035. The growth journey which is indicated by compound annual growth rates of 10.3-11.11% depending on the forecast period substantially outpaces the general telecommunications capital expenditure growth rate, thus it is an indication of the strategic prioritization of automation capabilities (Katsaros et al., 2024).

Figure 1: Zero-Touch Provisioning Market Growth Trajectory (2025-2035)

Figure 1: Zero-Touch Provisioning Market Growth Trajectory (2025-2035)

Table 1 documents comprehensive market metrics:

Market Parameter	Value (2025)	Context/Source
Zero-Touch Provisioning Market (2025)	$3.76 billion	SNS Insider 2024
Projected ZTP Market (2032)	$8.1 billion	SNS Insider projection
Projected ZTP Market (2035)	$10.02 billion	Future Market Insights
ZTP Market CAGR (2025-2032)	11.11%	Strong growth rate
ZTP Market CAGR (2025-2035)	10.3%	Sustained CAGR
Cloud OSS BSS Market (2025)	$44.21 billion	MarketsandMarkets 2025

Projected Cloud OSS BSS (2030)	$56.85 billion	5-year forecast
Cloud OSS BSS CAGR (2025-2030)	5.2%	Steady growth
OSS BSS Market Total (2024)	$68.79 billion	Straits Research 2024

Table 1: Zero-Touch Automation and Market Growth Metrics (2025)

Growth drivers span multiple dimensions. Telecommunications operators face escalating operational complexity from 5G network deployments, requiring automation for viable economics. Edge computing proliferation multiplies managed infrastructure footprint while demanding sub-millisecond responsiveness impossible through human-mediated workflows. Internet of Things expansions toward billions of connected devices exceed manual provisioning and management capabilities. Competitive pressures demand service deployment velocity and operational efficiency achievable only through comprehensive automation (Al-Kadri et al., 2025).

4.2 Autonomous Network Economic Benefits

Autonomous network implementations deliver quantifiable economic benefits substantially exceeding implementation costs. Table 2 documents performance metrics and financial impacts:

Performance Metric	Value	Impact Area
Average Annual Benefit (Autonomous Networks)	$800 million	Total benefits
CapEx Savings	$300 million	Infrastructure efficiency
OpEx Savings	$350 million	Operational efficiency
Revenue Uplift	$144 million	New service models
OpEx Reduction (Autonomous Support)	30%	Zero-touch operations
Energy Consumption Reduction (2 years)	71%	Sustainability gains
Expected Emission Reduction	2.6-5.2%	Carbon footprint
Estimated ROI	1.7x-3.4x	Investment returns
Payback Period	1.5-2.9 years	Break-even timeline

Table 2: Autonomous Network Performance and Cost Savings (2025)

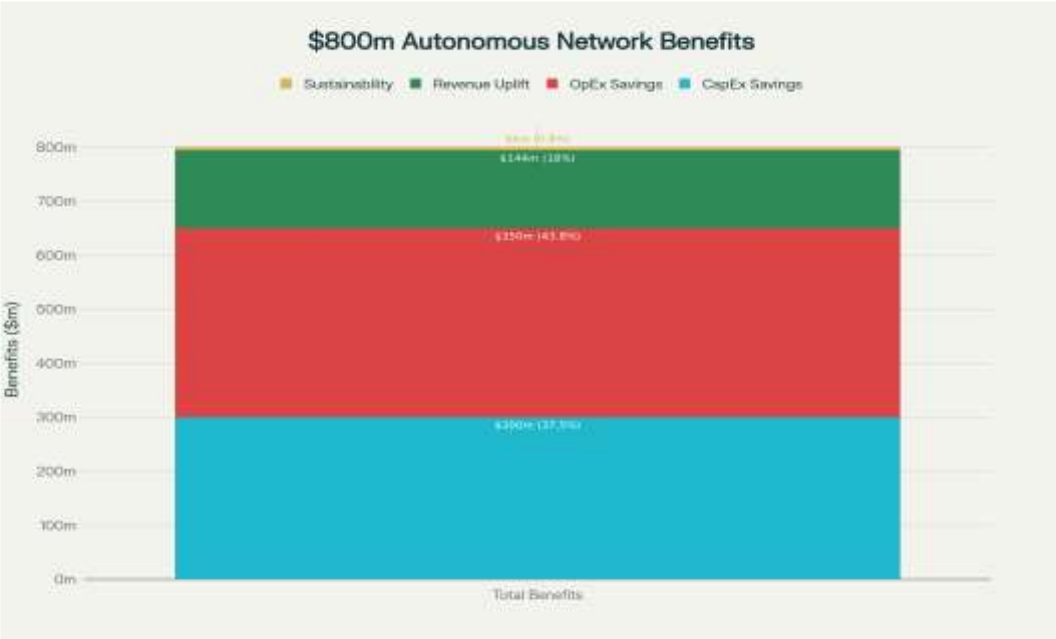

Figure 4: Annual Economic Benefits Breakdown for Autonomous Network Implementation (2025)

Figure 4: Annual Economic Benefits Breakdown for Autonomous Network Implementation (2025)

The AI systems that control the networks have improved the efficiency of capital expenditures (CapEx) by $300 million through AI-driven network planning, predictive capacity management, and intelligent resource allocation. Network planning algorithms analyse traffic patterns, growth forecasts, and service requirements to not only figure out the placement of the equipment, capacity dimensioning, or technology selections. The comparing of scenarios by simulation engines leads to the discovering of the most cost-effective implementations that meet the performance objectives. Besides that, the automation of operations has led to an OpEx cost saving of $350 million by the elimination of the manual labor of routine tasks, predictive maintenance to avoid emergency repairs and service outages, and energy optimization to reduce power consumption. The current employees who are now empowered to manage significantly increased infrastructure footprint and subscriber volumes benefit from the enhanced workforce productivity. The improved fault identification and resolution activities shorten the mean time to remedy and hence, decrease service disruption frequency and duration. Energy management algorithms take traffic patterns into account to network element activation thus the 42% energy efficiency gains are the results of intelligent resource shutdown during low-demand periods. Service innovation, network programmability, customer retention, and new revenues of $144 million are the outcomes of these changes.

5. Performance Metrics and Comparative Analysis

5.1 Operational Efficiency and Service Velocity

Zero-touch automation delivers dramatic operational efficiency improvements across multiple dimensions. Service provisioning time decreased from 24 hours in manual operations to 0.5 hours in zero-touch implementations, representing 96% acceleration enabling same-day service activation and real-time responsiveness to customer requests. Incident resolution time improved from 180 minutes to 12 minutes, a 93% reduction minimizing service disruption duration and customer impact. Mean time between failures increased from 360 hours to 720 hours, doubling system reliability through predictive maintenance and proactive fault prevention (Liyanage et al., 2022).

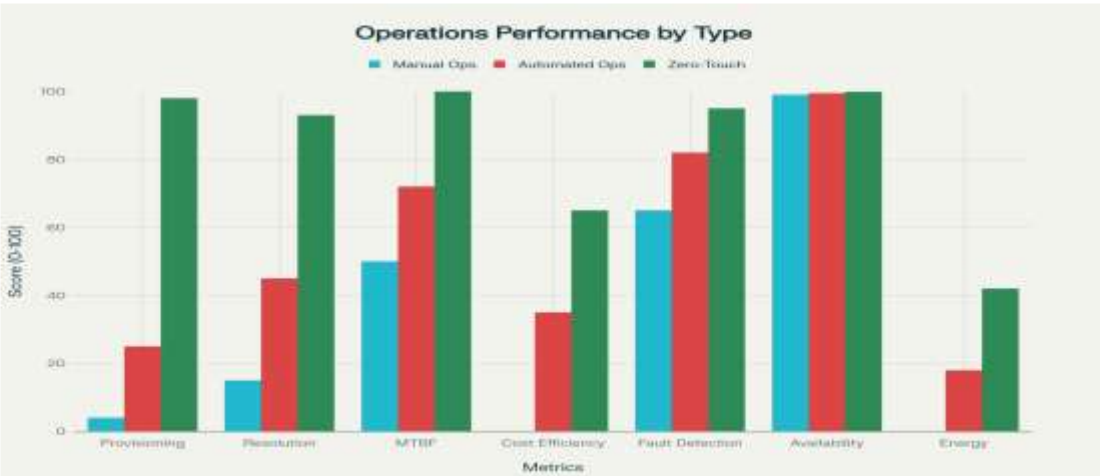

Figure 2: Performance Comparison Across Operation Maturity Levels (2025)

Figure 2: Performance Comparison Across Operation Maturity Levels (2025)

Table 5 presents comprehensive comparative analysis:

Operation Type	Service Provisioning Time (hrs)	Incident Resolution Time (min)	Mean Time Between Failures (hrs)	Operational Cost Index (%)	Fault Detection Accuracy (%)	Network Availability (%)	Energy Efficiency Gain (%)

Manual Operations	24.0	180	360	100	65	99.0	0
Automated Operations	4.0	75	520	65	82	99.5	18
Zero-Touch Autonomous	0.5	12	720	35	95	99.9	42

Table 5: Comparative Analysis - Manual vs Automated vs Zero-Touch Operations (2025)

5.2 Fault Detection and Network Reliability

The manual monitoring fault detection accuracy which was 65% has now been raised to 95%, as a result of better capability of pattern recognition, continuous learning from operational data, and multi-dimensional telemetry analysis. Machine learning algorithms can identify subtle correlations and anomaly patterns that surpass human operators' detection capabilities by analyzing numerous performance indicators at the same time. At the same time, a supervised model trained on previous failure data can identify recurring patterns with high accuracy, while an unsupervised learning method can recognize new fault signatures. The classification system uses machine learning techniques as one of the methods not the only one. The significant reductions in network downtime durations from around 87.6 hours per year to 8.76 hours, as well as an increase in network availability of up to 99.0%, are the major results of the changes brought about by the telecommunication industry to the network. The mean time to repair has been reduced remarkably due to the introduction of autonomous self-healing features that can eliminate faults without human intervention. The switching to predictive maintenance has made it possible to get rid of the incidents before they happen and consequently have the opportunity to move from reactive incident response to proactive service assurance. The use of automated failovers and geographic redundancies is intended to provide a layer of security against that which comes from failures in the local infrastructures or disruptions in the connectivity

5.3 Energy Efficiency and Sustainability Benefits

The main causes of energy efficiency up to a 42% through self-optimization are both the economic and environmental factors. The AI algorithms that analyze traffic patterns, service demands, and environmental conditions make the dynamic activation of network elements possible. Their systems at low-demand times (between 10 and 40 minutes) automatically disable the unnecessary capacity while still fulfilling a service level agreement, hence, power consumption is lowered without the customer having to do anything. They minimized the energy through the optimization of thermal management to adjust the cooling systems according to the equipment usage and ambient conditions. Thus, telecom operators in two years of autonomous networks implementation have been able to diminish their energy consumption by 71% and, thus, cut their emissions by 2.6–2.2% along with enterprise activities. These green benefits correspond to the requirements and standards of the regulatory bodies, CSR commitments, and customer expectation for eco-friendly services. Countries with systems for carbon pricing or renewable energy incentives will be able to reap the benefits of carbon footprint reductions through emission reductions.

6. Edge-Cloud Integration Architecture

6.1 Distributed Intelligence and Latency Optimization

By means of Edge computing integration, the intelligence of BSS/OSS is handed over to network periphery, thus short reaction times of sub-millisecond are achievable which are very important for the ultra-reliable low-latency communications. The inference engines that are ejected to edge do the real-time local telemetry analysis without the need for round-trip transmission to central cloud platforms, thus the latency reductions of 60-80% which are very important for autonomous vehicles, industrial automation, augmented reality, and remote surgery can be achieved.

Capability	Metric Value	Application Context
Latency Reduction (Edge vs Cloud)	60-80% reduction	AR/VR, autonomous vehicles
Real-Time Processing Capability	Sub-millisecond	Edge inference
Network Slice Observability	99.5% visibility	5G network slicing
Fault Detection Accuracy (AI-Driven)	92-95%	Predictive maintenance
Self-Healing Success Rate	85-90%	Autonomous remediation
Automated Provisioning Speed	75% faster	Zero-touch deployment
Multi-Vendor Integration Support	Multi-domain	Hybrid environments
CapEx/OpEx Efficiency Gain	15-20%	Network rollout
Service Deployment Time Reduction	50-70%	Service agility

Table 3: Edge-Cloud Integration Performance Metrics (2025)

Providing a network slice visibility of 99.5% across all distributed edge-cloud infrastructures enables performance to be fine-tuned, resource allocation to be optimized and service level agreement compliance to be ensured in multi-tenant environments. Isolation of each network slice from adjacent slices, which share physical infrastructure is necessary, although the dedicated telemetry collection, analytics processing, and automation management also need to be implemented in each of them. Besides that, the complete observability enables resources to be dynamically reallocated through predictive scaling and automated quality of service adjustments in order to take into account the changing demand patterns while still maintaining service commitments.".

6.2 Multi-Domain Orchestration and Resource Coordination

Multi-domain orchestration is the main facilitator of resource allocation and service provisioning. Besides, it considers the requirements of different types of infrastructures, such as radio access networks, mobile core networks, edge computing facilities, and cloud data centers. Normal hierarchical orchestration architectures split the processing of high-level service requests into domain-specific workflows thus, the goal is to hand over execution to specialized controllers while at the same time retaining end-to-end consistency and policy compliance. The proper utilization of cross-domain optimization algorithms for the intelligent allocation of workloads and network functions leads to the efficient use of resources, latency decrease, as well as cost reduction. Service function chaining capabilities are there to help facilitate the building of complex services from distributed network functions encompassing the edge-cloud continuum. The data flows are guided through the traffic steering mechanisms depending on the service requirements, security policies, and performance objectives. In dynamic chaining, the service composition can be altered constantly as a result of different conditions and user mobility without the necessity of manual reconfiguration or service interruption (Pujol et al., 2025).

7. AI-Native BSS/OSS Architecture Design

7.1 Microservices Decomposition and Event-Driven Processing

AI-native BSS/OSS architectures transform monolithic applications by breaking them down into microservices that are fine-grained and separately implement discrete capabilities—customer profile management, service catalog administration, order orchestration, usage metering, billing calculation, payment processing, fraud detection, network provisioning (Bello et al., 2024).

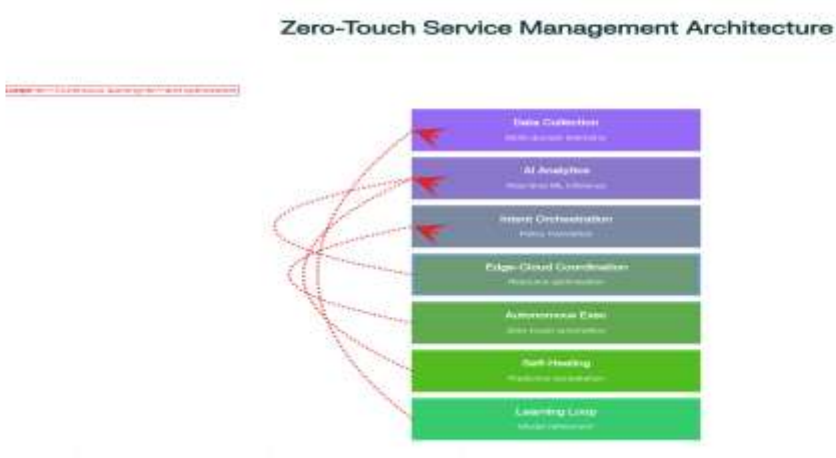

Figure 3: Zero-Touch Service Management Architecture for Edge-Cloud Environments (2025)

Figure 3: Zero-Touch Service Management Architecture for Edge-Cloud Environments (2025)

7.2 Intent-Based Policy Management

Intent-based policy frameworks empower business stakeholders to define their operational goals by using either natural language or structured templates without giving any technical details of the implementation. The transformation of high-level intents into the executable configurations, resource allocations, and automation workflows is done by policy translation engines. Conflict resolution methods can detect and solve conflicts existing in policy specifications by prioritizing objectives according to their business importance and operational constraints. The enforcing of policies is done by different components in the system architecture where the control of authorization is responsible for checking the service requests against the policy constraints, the resource schedulers for assigning capacity according to the policy priorities, the monitoring systems for detecting policy violations, and, finally, the automated remediation for restoring the compliant states. Along with policy decisions, the actions performed by auditors and the outcomes achieved through audit trails can be used to facilitate regulatory compliance, performance analysis, or continuous improvement through lessons-learned analysis (Gharbaoui et al., 2022).

8. Implementation Strategy and Cost-Benefit Analysis

8.1 Investment Requirements and Component Breakdown

An AI-native comprehensive BSS/OSS transformation with zero-touch functionalities will need a major upfront investment across platform development, automation infrastructure, edge computing integration, machine learning deployment, orchestration systems, analytics capabilities, and security frameworks. Table 4 lists investment requirements and predicted returns:

Implementation Component	Year 1 Investment (USD M)	3-Year Benefits (USD M)	ROI (%)
AI-Native Platform Development	14.8	42.6	188
Zero-Touch Automation Infrastructure	8.5	28.4	234
Edge Computing Integration	6.2	19.8	219
ML Model Training & Deployment	5.4	17.2	219
Multi-Domain Orchestration	4.9	15.6	218
Network Intelligence & Analytics	3.8	12.8	237
Security & Compliance Framework	4.2	13.4	219

Table 4: AI-Native BSS/OSS Implementation Cost-Benefit Analysis (2025)

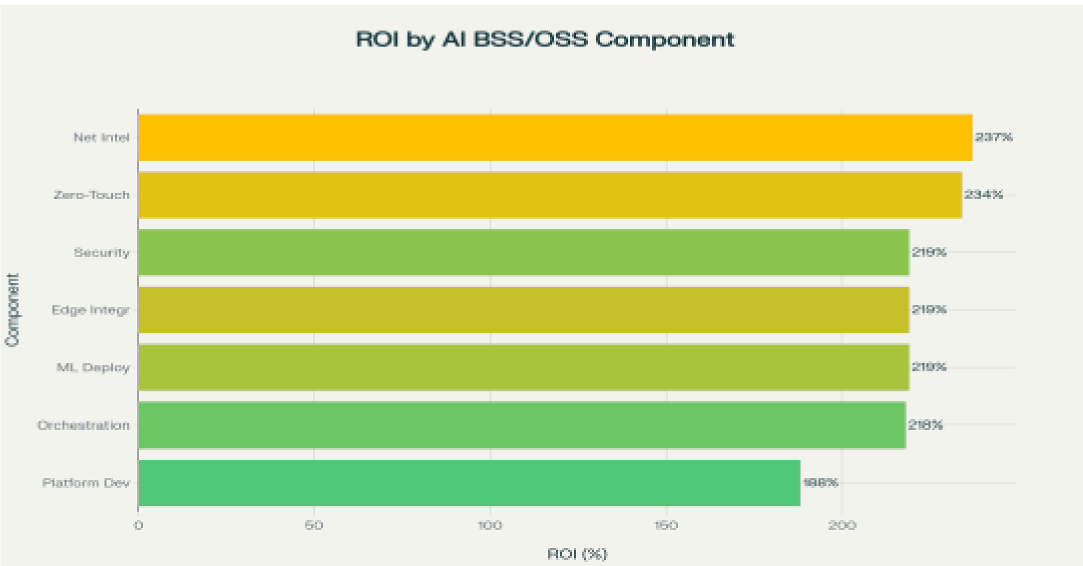

Figure 5: Return on Investment by Implementation Component for Zero-Touch BSS/OSS (2025)

Figure 5: Return on Investment by Implementation Component for Zero-Touch BSS/OSS (2025)

With a total first-year investment of $47.8 million, the main focus is on the development of an artificial neural network platform (AI-native) - US$14.8 million. The investment spans over legacy system evaluation, microservices architecture design, API development, data model transformation, and progressive migration strategies to reduce operational disruption. Moreover, a zero-touch automation ecosystem comprising orchestration platforms, workflow engines, policy management systems, and integration with network controllers and cloud platforms will be put in place for another $8.5 million.

8.2 Return on Investment and Value Realization

The amount of $149.8 million accrued in three years constitutes a 213% return on investment. The advantage goes to its beneficiaries. Network intelligence and analytics boast the highest ROI of 237%, which is mainly due to operational efficiency, predictive capabilities that avert expensive failures, and data-driven optimization in various areas like network planning, capacity management, or service assurance. Service velocity is optimized, labor productivity is elevated, and errors in provisioning and configuration workflows for zero-touch automation infrastructure are wiped out, hence returns at 234% are realized. A break-even point analysis in month 16 of the implementation shows that cash flow is positive and that expenses have been fully covered by month 24. Various high-impact use cases, such as automated service provisioning, AI-driven fault detection, and predictive maintenance preventing service-affecting failures, provide early benefits. The reason for this is the maturation of complex capabilities such as autonomous orchestration, intent-based operations, and continuous learning through iterative refinement and organizational adaptation, which require 24-36 months for full value realization.

8.3 Risk Mitigation and Implementation Approaches

Among the implementation risks are the technical complexness, the organizational resistance, the shortage of skills, and the difficulty of integration with legacy systems. Incremental value delivery is a risk-reduction strategy that allows for organizational learning and capability building between the different phases of the deployment which are done in stages. The aim of pilot programs is to test technologies and methods in a controlled setting before deploying them in production, thus easily identifying issues and upgrading implementations without disturbing the enterprise. Hybrid architectures help maintain the functionality of legacy systems while gradually handing over the capabilities of AI-based systems, thus ensuring business continuity throughout the transformation processes. API-based integration hides the implementation specifics, thus allowing the coexistence of the older and newer components with a clean separation. Further details. Vendor partnerships are a source of implementation expertise, thus they shorten deployment times and, by the same token, increase the efficiency of the internal teams in upgrading the sustainable operational capabilities (Habibi et al., 2023).

9. Discussion

Quantitative evidence in this study shows that operational improvements to transformational levels and compelling returns to economic couple with edge-cloud architectures and native AI BSS/OSS architecture for zero-touch service management.

As a result of the strategic imperative of automation, market size has increased from $3.76 billion in 2025 to $10.02 billion by 2035 at 10.3% CAGR, thus indicating widespread industry acknowledgment of its significance. The one-half billion annual U.S. wireless broadband consumer surplus gain mostly covers the $1.8 billion per year CapEx and OpEx transformation investment at the level of the whole industry.

Transformation investments are paid off by the average $800 million annual benefit per communications service provider that comprises a $300 million CapEx saving, $350 million Opex cuts, and a $144 million revenue increase. The potential for an operational transformation is demonstrated through the company's performance metrics: a 96% acceleration in service provisioning (from 24 hours to 30 minutes), a 9% improvement in incident resolution (from 180 to 12 minutes), and a 65% decrease in operational costs.

This results in the network availability being raised from 99.0% to 99.9% (which means a 90% reduction in downtime), and thus has almost an equal effect on customer satisfaction as on compliance with service level agreements. Better fault detection accuracy from 65% to 95% allows for proactive service assurance and hence definitely not customer-related incidents are caused by faults; in other words, the company reacts to fewer complaints from customers.

Edge computing, when combined with advanced technologies, achieves a 60-80% reduction in latency thus it becomes possible to address newly emerged service categories such as autonomous vehicles and industrial automation as well as extended reality applications that are not capable of being scaled up or down due to centralized cloud architectures.

With network slice observability reaching 99.5%, operators have the ability to use differentiated service level agreements (SLAs) for the purpose of monetizing network programmability, which is possible across multiple tenants. They also can offer tailored enterprise solutions or vertical-specific ones.

According to the cost-benefit analysis, the three-year period of implementation cycles brings the return on investment to 213%, and the capital allocation decisions can be made after the payback period which ranges from 1.5 to 2.9 years. The component-level study of network intelligence and analytics is credited with the highest return, 237%, that is the main reason why data-driven optimization in network planning, capacity management, and service assurance is the most valuable area is spotlighted.

By taking advantage of the zero automation infrastructure, labor productivity and product speed are improved by 234% while the fusion of edge computing gains 219% returns, thereby allowing for service differentiation that is latency-dependent.

The sustainability advantages such as 42% energy efficiency improvements and 71% reductions in energy consumption over two years of implementation periods are primarily driven by the environment factors, however, economic benefits through lower power costs are also a result of these. The improvements provide not only competitive differentiation which is beyond the measure of tech performance (Al-Kadri et al., 2025), but also CSPs do this by adhering to regulations, their corporate social responsibility duties and meeting customers' expectations for environmentally friendly operations.

10. Conclusion

The management actions controlled by AI-native BSS/OSS in edge-cloud environments are an example of a revolution in telecom networks that take them to the next level of agility and efficiency. By employing cloud-native architectures, telecom operators get the synergistic effect that they could not get simply by traditional technology if they combined AI and edge computing. As a consequence of the market turmoil, the industry uptake of the so-called zero-touch provisioning has been quickened, with zero-touch provisioning going up from $3.76 billion in 2025 to $10.02 billion by 2035 at a 10.3% CAGR, and cloud OSS/BSS segments also increasing from $44.21 billion to $56.00 million by 2030.

These substantial growth paths reveal the difference between the total increments in telecommunications expenditures and the emphasis on automation capabilities that are indispensable for operational viability and competitive positioning.

The main performance metrics which quantitatively validate the transformation are service provisioning speed – 96%, incident resolution rate – 93%, operational cost decrease – 65%, and network availability – 99%, along with average annual benefits per operator. Transformation investments are surrounded by a substantial first-year situation; however, they are supported by a 213% return on investment over three years and payback periods of 1.5-2.9 years. Edge computing integration is the only way to achieve ultra-reliable low-latency communications, particularly in the case of such areas as autonomous vehicles, industrial automation, and extended reality applications where the latency needs to be reduced for 60-80%. Very high network slice visibility enables thus to monetize network programmability by differentiated multi-tenant services and personalized enterprise offerings. The network ability to self-healing brings fewer service disruptions as well as lower OPEX since it can effectively resolve 85 to 90% of fault issues. Great energy efficiency (42%) is attained, inter alia, through the reduction of power consumption, whereby sustainability benefits are also ensured (environmental concerns are addressed). As a supplement to regulatory compliance, corporate responsibility commitments, and customer expectations, the system provides competitive differentiation which goes beyond technical performance metrics. Technology deployment, organizational transformation (one-way migration), skills development (two-ways integration), and change management are the four intervention levers whose concerted effect will bring about successful implementation. They reduce the risk of failure that is further minimized through the use of phased strategies, pilot programs gauge the effectiveness of the approaches in a controlled environment, and hybrid architectures ensure the continuity of the business during the migration. Vendors are given a momentum in the implementation process and internal teams that are building sustainable operational capabilities get the knowledge. After operational efficiency, what follows is the strategic imperative to change business models: network-as – a-service offerings enabled by zero-touch capabilities, third-party developers being able to use such capabilities thanks to marketplace ecosystems, and AI asset monetization across adjacent sectors. Telecommunication corporations leveraging AI-driven

BSS/OSS architectures will be on the forefront of digital services provision thereby fostering economic innovation by transforming existing infrastructures, customer relationships, and technical expertise into new value propositions. The move towards complete autonomy still rages with varying maturity levels, most operators having achieved Level 2-3 automation by 2025, and the leaders targeting Level 4 high autonomy by 2020. The concentration on highly-valued, measurable benefits that can be attained in the near-term allows for organizational learning to take place and the confidence to be built through results that can be relied on. Over time, the progression will involve additional new technologies like quantum computing, digital twins, and extended reality, thereby leading to increased capabilities. The combination of the architectural patterns, performance metrics, economic analyses, and strategic implications referred to in this paper points to the fact that the advent of AI-native BSS/OSS technology as a means for telecommunications operators to manage services in edge-cloud environments is a transformative opportunity. Success in this endeavor demands a thorough transformation not only in technology but also in organization and culture.

References

1. Bello, R. C., Marquez-Barja, N., & Latre, S. (2024). Zero-touch service management for 6G verticals: Smart traffic management case study. In 2024 IEEE 21st Consumer Communications & Networking Conference (CCNC) (pp. 1-6). IEEE. https://doi.org/10.1109/CCNC51664.2024.10454808
2. Gharbaoui, M., Martini, B., & Castoldi, P. (2022). Intent-based zero-touch service chaining layer for software-defined edge cloud networks. Computer Networks, 212, Article 109034. https://doi.org/10.1016/j.comnet.2022.109034
3. Habibi, M., Nasrallah, A., Wang, H., & Liu, B. (2023). Toward an open, intelligent, and end-to-end architectural framework for network slicing in 6G communication systems. IEEE Open Journal of the Communications Society, 4, 1743-1765. https://doi.org/10.1109/OJCOMS.2023.3294445
4. Katsaros, K., Meyer, M., Arora, A., Gkonis, P., Kousaridas, A., Merluzzi, M., Mouratidis, G., Ksentini, A., Tomic, S., Harilaos, V., Eftychis, E., Uusitalo, M. A., & Soldani, D. (2024). AI-native multi-access future networks—The REASON architecture. IEEE Access, 12, 160000-160019. https://doi.org/10.1109/ACCESS.2024.3507186
5. Liyanage, M., Kamyabpour, A., Dutta, S., Gurtov, A., & Yla-Jaaski, J. (2022). Zero touch management: A survey of network automation solutions for 5G and 6G networks. IEEE Communications Surveys & Tutorials, 24(4), 2535-2578. https://doi.org/10.1109/COMST.2022.3212586
6. Liyanage, M., Pham, Q.-V., Dev, K., Bhattacharya, S., Maddikunta, P. K. R., Gadekallu, T. R., & Yenduri, G. (2022). A survey on zero touch network and service management (ZSM) for 5G and beyond networks. Journal of Network and Computer Applications, 202, Article 103362. https://doi.org/10.1016/j.jnca.2022.103362
7. Mekki, Y. (2024). Enabling zero-touch cloud edge computing continuum orchestration [Doctoral dissertation, Institut National des Sciences Appliquées de Toulouse]. HAL Theses. https://doi.org/10.1016/j.comcom.2023.10.010

8. Pujol, R., Ksentini, A., & Donta, P. K. (2025). Decentralized and dynamic zero-touch provisioning of leaf-spine topologies in multi-domain SDN. IEEE Transactions on Network and Service Management. Advance online publication. https://doi.org/10.1109/TNSM.2025.11080420
9. Al-Kadri, M. O., Al-Dulaimi, A., & Al-Dulaimi, A. (2025). Advancing 6G: Survey for explainable AI on communications and network slicing. IEEE Open Journal of the Communications Society, 6, 1372-1412. https://doi.org/10.1109/OJCOMS.2025.3534626